THEMIS SIGNAL ANALYSIS STATISTICS RESEARCH PROGRAM

CORRELATION BETWEEN TWO HOTELLING'S T^2

by

A. M. Kshirsagar* and John C. Young

Technical Report No. 79
Department of Statistics THEMIS Contract

Department of Statistics
Southern Methodist University
Dallas, Texas 75222

9

THEMIS SIGNAL ANALYSIS STATISTICS RESEARCH PROGRAM

CORRELATION BETWEEN TWO HOTELLING'S T^2

by

A. M. Kshirsagar and John C. Young

Technical Report No. 79
Department of Statistics THEMIS Contract

August 1, 1970

Research sponsored by the Office of Naval Research
Contract N00014-68-A-0515
Project NR 042-260

DEPARTMENT OF STATISTICS
Southern Methodist University

CORRELATION BETWEEN TWO HOTELLING'S T^2

by

A. M. Kshirsagar* and John C. Young

Southern Methodist University
Dallas, Texas 75222

1. Introduction:

Let B be a p×p symmetric matrix having the Wishart distribution

(1.1) $$W_p(B|I|f)dB = C_{pf}|B|^{(f-p-1)/2}\, e^{-1/2\ \mathrm{tr}B}\, dB\ ,$$

where

(1.2) $$C_{pf}^{-1} = 2^{fp/2}\, \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(\frac{f+1-i}{2}\right)\ ,$$

and dB stands for the product of the differentials of the p(p+1)/2 distinct elements of B. Let $\underline{x}$ and $\underline{y}$ be two vector variables of p components, distributed independently of B, and also independently of each other, as

(1.3) $$\frac{1}{(2\pi)^{p/2}}\, e^{-1/2\ \underline{x}'\underline{x}}\, d\underline{x}\ ,$$

and

(1.4) $$\frac{1}{(2\pi)^{p/2}}\, e^{-1/2\ \underline{y}'\underline{y}}\, d\underline{y}$$

respectively. While considering the problem of multivariate statistical outliers, Wilks (1963) used statistics of the type,

*Research supported by ONR Contract N00014-68-A-0515, Department of Statistics THEMIS Project.

(1.5) $r = |B + \underline{yy}'|/|B + \underline{xx}' + \underline{yy}'|$, $s = |B + \underline{xx}'|/|B + \underline{xx}' + \underline{yy}'|$.

He has remarked that the exact distribution (joint) of r and s is complicated and has given the expected values, variances and covariance of r and s. Unfortunately, his expressions for the variance and covariance are in error. The purpose of this note is to derive the exact joint distribution of r and s and to give correct expressions for the moments.

2. Joint distribution:

In the joint distribution of B, $\underline{x}$ and $\underline{y}$, make the following transformation

$$(2.1) \qquad \begin{aligned} A &= B + \underline{x}\,\underline{x}' + \underline{y}\,\underline{y}' , \\ \underline{u} &= A^{-1/2}\,\underline{x} , \\ \underline{v} &= A^{-1/2}\,\underline{y} , \end{aligned}$$

where $A^{-1/2}$ is any matrix such that $A^{-1/2} \cdot A^{-1/2} = A^{-1}$. The Jacobian of transformation from B to A is 1 and that from $\underline{x}$ to $\underline{u}$ or $\underline{y}$ to $\underline{v}$ is $|A|^{1/2}$ and hence, the joint distribution of A, $\underline{u}$ and $\underline{v}$ comes out as

$$(2.2) \qquad \frac{C_{pf}}{(2\pi)^p} |A|^{\frac{(f+2)-p-1}{2}} e^{-1/2 \,\mathrm{tr}A} \cdot |I - \underline{uu}' - \underline{vv}'|^{\frac{f-p-1}{2}} \, dA d\underline{u} d\underline{v} ,$$

as $|B| = |A - A^{1/2}\underline{uu}'A^{1/2} - A^{1/2}\underline{vv}'A^{1/2}| = |A||I - \underline{uu}' - \underline{vv}'|$.

This shows that A has a Wishart distribution of f+2 degrees of freedom and is independent of $\underline{u}$ and $\underline{v}$. Splitting the constant suitably, the joint distribution of $\underline{u}$ and $\underline{v}$ is

$$(2.3) \qquad \frac{\Gamma(f+1)}{(2\pi)^p \Gamma(f-p+1)} |I - \underline{uu}' - \underline{vv}'|^{\frac{f-p-1}{2}} \, d\underline{u} d\underline{v} .$$

Observe that the statistics r, s of Wilks are given by

$$(2.4) \qquad r = \frac{|B+\underline{yy}'|}{|B+\underline{xx}'+\underline{yy}'|} = \frac{|A-\underline{xx}'|}{|A|} = |I - \underline{uu}'| = 1 - \underline{u}'\underline{u} ,$$

and

$$(2.5) \qquad s = \frac{|B+xx'|}{|B+xx'+yy'|} = 1 - \underline{\nu}'\underline{\nu} \quad .$$

Also observe that in (2.3)

$$(2.6) \qquad \begin{aligned} |I - \underline{uu}' - \underline{\nu\nu}'| &= (1 - \underline{u}'\underline{u})(1 - \underline{\nu}'\underline{\nu}) - (\underline{u}'\underline{\nu})^2 \\ &= rs - (\underline{u}'\underline{\nu})^2 \quad . \end{aligned}$$

In (2.3), transform from $\underline{\nu}$ to $\underline{w} = [w_1, w_2, \ldots, w_p]'$, by an orthogonal transformation

$$(2.7) \qquad \underline{w} = L\underline{\nu} \ ,$$

where

L is a p×p orthogonal matrix, whose last row is $\underline{u}'/\sqrt{\underline{u}'\underline{u}}$. The Jacobian of this transformation is $|L| = 1$ and $\underline{\nu}'\underline{\nu} = \underline{w}'\underline{w} = 1-s$. Also

$$(2.8) \qquad u'\nu = \underline{u}'L'L\underline{\nu} = [0 \ldots 0, \sqrt{u'u}\,]\underline{w} = \sqrt{1-r} \cdot w_p \quad .$$

The joint distribution of $\underline{u}$ and $\underline{w}$ is, therefore,

$$(2.9) \qquad \frac{\Gamma(f+1)}{(2\pi)^p \Gamma(f-p+1)} \left\{ rs - (1-r)w_p^2 \right\}^{\frac{f-p-1}{2}} d\underline{u}d\underline{w} \quad .$$

From $\underline{u}$, transform to $r = 1 - u'u$ and p-1 other variables $\phi_1, \phi_2, \ldots, \phi_{p-1}$ by

$$(2.10) \qquad \begin{aligned} u_1 &= (1-r)^{1/2}\cos\phi_1 \cos\phi_2 \ldots \cos\phi_{p-1} \ , \\ u_j &= (1-r)^{1/2}\cos\phi_1 \cos\phi_2 \ldots \cos\phi_{p-j} \sin\phi_{p-j+1} \\ & \qquad (j=2, 3, \ldots p) \end{aligned}$$

Similarly, transform from $\underline{w}$ to $s = 1 - \underline{w}'\underline{w}$ and p-1 other variables $\theta_1, \theta_2, \ldots, \theta_{p-1}$ by

$$(2.11) \qquad \begin{aligned} \nu_1 &= (1-s)^{1/2}\cos\theta_1 \cos\theta_2 \ldots \cos\theta_{p-1} \ , \\ \nu_j &= (1-s)^{1/2}\cos\theta_1 \cos\theta_2 \ldots \cos\theta_{p-j} \sin\theta_{p-j+1} \quad . \\ & \qquad (j=2, 3, \ldots, p) \end{aligned}$$

The Jacobian of transformation from $\underline{u}$ to $r, \phi_1, \dots \phi_{p-1}$ is

$$\frac{1}{2}(1-r)^{\frac{1}{2}p-1} \prod_{i=1}^{p-2} \cos^{p-i-1}\phi_i$$

and a similar expression in s and θ_i for the Jacobian of transformation from $\underline{w}$ to s and the θ's. Now θ_{p-1} and ϕ_{p-1} vary from 0 to 2π, the other θ's and ϕ's vary from $-\pi/2$ to $\pi/2$ while r and s vary from 0 to 1. Integrating out all the ϕ's and all θ's except θ_1, we obtain the joint distribution of r, s and θ_1 as

$$(2.12) \qquad \frac{\Gamma(f+1)}{4\pi\Gamma(p-1)\Gamma(f-p+1)} \left\{rs - (1-r)(1-s)\sin^2\theta\right\}^{\frac{f-p-1}{2}} \cos^{p-2}\theta \; dr ds d\theta$$

where θ_1 is replaced by θ.

The joint distribution of r,s alone can now be obtained by integrating out θ but this does not seem to yield a manageable expression, as the bracket in (2.12) will have to be expanded in a series.

3. Moments of r,s .

Only the product moment of r and s is difficult to obtain. The mean and variance of r (or s) can be very easily obtained from the marginal distribution of r, which is related to the well-known Hotelling's T^2 by $r = \dfrac{1}{1+\left(\dfrac{T^2}{f+1}\right)}$. In the joint distribution of $\underline{u}$ and $\underline{v}$, given by (2.3), if we transform to $\underline{z} = [z_1, \dots, z_p]'$ from $\underline{v}$ by

$$(3.1) \qquad \underline{v} = (I - \underline{uu}')^{1/2}\underline{z} \quad ,$$

we shall find that $\underline{u}$ and $\underline{z}$ are independently distributed as

$$(3.2) \qquad K(\underline{u}|f)d\underline{u} = \frac{f}{\pi^{p/2}(f-p)} \cdot \frac{\Gamma(f/2)}{\Pi\frac{}{2}(f-p)} |I - \underline{uu}'|^{\frac{f-p}{2}} d\underline{u}$$

(3.3) and $K(\underline{z}|f-1)d\underline{z}$, respectively.

From (3.2), one can easily show that

$$(3.4)\qquad E(r^h) = E(1-\underline{u}'\underline{u})^h = E|I - uu'|^h$$

$$= \frac{f(f+2h-p)}{(f-p)(f+2h)} \cdot \frac{\Gamma\left(\frac{f-p}{2}+h\right)\Gamma\left(\frac{f}{2}\right)}{\Gamma\left(\frac{f}{2}+h\right)\Gamma\left(\frac{f-p}{2}\right)}$$

This will also be the h^{th} moment of s by symmetry. This leads to

$$(3.5)\qquad E(r) = \frac{f-p+2}{f+2}, \quad V(r) = \frac{2p(f-p+2)}{(f+2)^2(f+4)} .$$

Now $\mathrm{Cov}(r,s) = E\{(1-\underline{u}'\underline{u})(1-\underline{v}'\underline{v})\} - E(r)E(s)$

$$= E\{(1-\underline{u}'u)[1-\underline{z}(I-\underline{uu}')\underline{z}]\} - \{E(r)\}^2 \quad \text{by (3.1)}$$

$$= E(r) - E\{r[\underline{z}'\underline{z}-(\underline{z}'\underline{u})^2]\} - \{E(r)\}^2$$

$$(3.6)\qquad = E(r) - E(r)E(z'z) + E\{r(z'u)^2\} - \{E(r)\}^2 ,$$

as $\underline{z}$ and r are independent. Since $\underline{z}$ has the same distribution as $\underline{u}$ with f changed f-1 ,

$$E(z'z) = 1 - E(1-\underline{u}'\underline{u}) \text{ with f replaced by f-1}$$

$$(3.7)\qquad = \frac{p}{f+1}$$

Hence (3.6) reduces to

$$(3.8)\qquad \mathrm{Cov}(r,s) = \frac{-p(f-p+2)}{(f+1)(f+2)^2} + E\{r(z'u)^2\} .$$

Now

$$(3.9)\qquad E\{r(\underline{z}'\underline{u})^2\} = \int (1-u'u)(\underline{z}'\underline{u})^2 K(\underline{u}|f)K(\underline{z}|f-1)d\underline{u}d\underline{z}$$

where the integration is over the range of values of $\underline{u}$ and $\underline{z}$ such that $\underline{u}'\underline{u} \le 1$, $\underline{z}'\underline{z} \le 1$. Transform from $\underline{z}$ to $\underline{\xi} = [\xi_1, \ldots, \xi_p]$ by the transformation

$$\xi = L\underline{z} ,$$

where L is already defined to be a p×p orthogonal matrix, whose last row is $\underline{u}'/\sqrt{\underline{u}'\underline{u}}$. Then,

$$\underline{z}'\underline{u} = \underline{z}'L'L\underline{u} = \xi'L\underline{u} = \xi_p\sqrt{\underline{u}'\underline{u}} = (1-r)^{1/2}\xi_p .$$

Hence (3.9) reduces to

$$(3.10) \qquad \int r(1-r)K(\underline{u}|f)d\underline{u} \cdot \int \zeta_p^2 \, K(\underline{\zeta}|f-1)d\underline{\zeta} = E(r-r^2) \cdot \frac{1}{p} E(\underline{\zeta}'\underline{\zeta}) \, ,$$ due

to symmetry of the distribution of $\underline{\zeta}$. Now $\underline{\zeta}$ has the same distribution as $\underline{u}$ with f replaced by $f-1$ and hence finally, (3.10) reduces to

$$\frac{p(f-p+2)}{(f+4)(f+2)} \cdot \frac{1}{f+1} \quad .$$

The covariance between r and s, therefore, is (from (3.5))

$$(3.11) \qquad \frac{-2p(f-p+2)}{(f+1)(f+2)^2(f+4)} \quad .$$

Remarks:

Wilks considers a sample of size n and has a Wishart matrix based on n-1 degrees of freedom as deviations are from the sample means. He then removes two observations as outliers and thus his (n-1)-2 corresponds to our f. His E(r) agrees with our result, with this correspondence but the other moments are in error.

Reference

Wilks, S. S. (1963). "Multivariate Statistical Outliers," Sankhyā, Vol. 25, p. 407-426.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author): SOUTHERN METHODIST UNIVERSITY	2a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	
	2b. GROUP: UNCLASSIFIED	
3. REPORT TITLE: Correlation between two Hotelling's T^2		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates): Technical Report		
5. AUTHOR(S) (First name, middle initial, last name): A. M. Kshirsagar; John C. Young		
6. REPORT DATE: August 1, 1970	7a. TOTAL NO. OF PAGES: 6	7b. NO. OF REFS: 1
8a. CONTRACT OR GRANT NO.: N00014-68-A-0515	9a. ORIGINATOR'S REPORT NUMBER(S): 79	
b. PROJECT NO.: NR042-260		
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.		
10. DISTRIBUTION STATEMENT: This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.		
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY: Office of Naval Research	

13. ABSTRACT

If B is a Wishart matrix and $\underline{x}$, $\underline{y}$ are two vectors of p components each having a multinormal distribution and if all these quantities are independently distributed, the joint distribution of the two statistics

$$r = \frac{|B + \underline{y}\underline{y}'|}{|B + \underline{x}\underline{x}' + \underline{y}\underline{y}'|} \quad \text{and} \quad s = \frac{|B + \underline{x}\underline{x}'|}{|B + \underline{x}\underline{x}' + \underline{y}\underline{y}'|}$$

is derived in this paper. The correlation between r and s is also obtained. r and s are related to Hotelling's T^2 and are useful in problems of testing multivariate outliers.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

UNCLASSIFIED

Security Classification